STRUCTURAL STUDIES OF THE MAJOR GLYCOPROTEIN IN PREPARATIONS WITH CARCINOEMBRYONIC ANTIGEN ACTIVITY*

By WILLIAM D. TERRY, PIERRE A. HENKART, JOHN E. COLIGAN, AND CHARLES W. TODD

(*From the Immunology Branch, National Cancer Institute, Bethesda, Maryland 20014, and the Department of Immunology, City of Hope National Medical Center, Duarte, California 91010*)

(Received for publication 28 April 1972)

Previous investigations have shown that carcinomas of the human colon contain a tumor-associated antigen (1). Antisera prepared against extracts of colon carcinoma demonstrated an antigen(s) present in all colon carcinomas, but absent from, or present at very low levels in, corresponding normal colonic tissue (2–4). The antigen was also identified in embryonic and fetal gastrointestinal tissues, and because of this, the material was designated the carcinoembryonic antigen (CEA)[1] of the human digestive system (2).

Preliminary characterization of the antigenic material extracted and partially purified from tumor tissue indicated that it was a glycoprotein (5). It seems important to characterize these extracts further, and for this reason, chemical studies of preparations with CEA activity are being carried out. This report describes preliminary results of amino acid sequence studies of the only protein in these preparations with an unblocked amino-terminal amino acid.

Materials and Methods

Colon cancer metastases were obtained from the livers of patients Be and Na 1 at autopsy and frozen at −10°C until used. A complete description of the isolation procedure used is presented elsewhere (3). In brief, tumor tissue was homogenized in a Waring blender (Waring Products Div., New Hartford, Conn.) and mixed with an equal volume of 2 M perchloric acid. The precipitate was removed by centrifugation, and the supernatant chromatographed on Sepharose 4B (Pharmacia Fine Chemicals, Inc., Piscataway, N. J.). Fractions containing CEA activity, as detected by a double antibody assay (6), were pooled and rechromatographed on a column of Sephadex G-200 (Pharmacia). Active fractions were pooled and lyophilized.

The Na 1 preparation was further purified by chromatography on epichlorohydrin triethanolamine cellulose brought to a pH of 6.5 and equilibrated with 0.01 M NaCl. Samples were applied in 0.01 M NaCl and eluted using an exponential gradient with a final concentration of 0.2 M NaCl. Multiple peaks were obtained and all contained CEA antigenic activity.

* This work was supported in part by grant No. 12631 from the National Cancer Institute.

[1] *Abbreviations used in this paper*: CEA, carcinoembryonic antigen; SDS, sodium dodecyl sulfate.

One portion of the column effluent designated Na 1–3 was chosen for further chemical characterization.

RESULTS AND DISCUSSION

Isolated Be material was studied by disc gel electrophoresis in 4% acrylamide gels with or without urea at pH 8.7 (7), and sodium dodecyl sulfate (SDS) gels containing 3.3% acrylamide (8). Gels stained for periodate oxidizable sugars (9) revealed a single diffuse band, with no discrete minor bands (Fig. 1). Gels containing 30–50 μg of CEA showed very slight diffuse staining with Coomassie brilliant blue in the region of the periodic acid-Schiff stain. No staining was detectable on these gels with amido schwarz or aniline blue black. The ineffec-

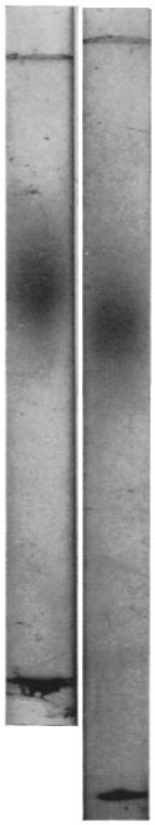

FIG. 1. Acrylamide gel electrophoresis of Be CEA preparations. The gel at the left contained 8 M urea and was run at pH 8.9, while the gel at the right contained 0.1% SDS. For both gels the sample was 60 μg of CEA and contained 1% β-mercaptoethanol. The bromthymol blue tracker dye was marked with India ink (bottom) before staining for periodate oxidizable sugars.

tiveness of protein staining of CEA is unexplained; gels containing other proteins stained as expected. These findings suggest that this CEA preparation contains one or more glycoproteins of similar size and charge characteristics and does not contain detectable protein components which, if present, should appear as separate sharp bands in the gels stained for protein.

The amino acid composition of the Be preparation (Table I) is unremarkable except for the absence of methionine. It should be noted that the composition is quite different from that previously reported for CEA (5) but is very similar to the composition we have obtained for CEA preparations from five other tumors. The amino acid composition of Na 1–3 is similar to that of Be, but there are distinct differences which may reflect the additional purification of Na 1–3 (Table I). Protein accounts for 25–50% by weight of the various CEA preparations. The reasons for this wide range in apparent protein content are not clear at the present time.

Amino acid sequence analysis of the glycoprotein preparations were performed using an automatic protein sequencer (Beckman 890; Beckman Instruments, Inc., Fullerton, Calif.). Phenylthiohydantoin amino acids were identified on a gas chromatograph (Varian 1840; Varian Aerograph, Walnut Creek, Calif.) (10), and by conversion to the free amino acid by hydrolysis with subsequent identification of the amino acid on the amino acid analyzer (11). The results of sequence analysis are shown in Fig. 2. 16 of the first 19 residues were identified in the Be preparation. Be was available only in limited quantities, and the amino acids at positions 1, 7, and 15 have not been definitely estab-

TABLE I

Amino Acid Composition of Two CEA Preparations

Amino acid	Be	Na 1-3
	Moles amino acid/10^5 g protein	
Lysine	23.5	28.6
Histidine	17.2	17.3
Arginine	33.9	33.5
Aspartic acid	142.0	124.8
Threonine	77.6	77.5
Serine	90.2	91.8
Glutamic acid	94.5	92.6
Proline	103.8	100.0
Glycine	49.0	48.6
Alanine	54.1	54.7
Valine	65.6	58.4
Methionine	<1.0	<1.0
Isoleucine	45.6	41.5
Leucine	82.3	77.5
Tyrosine	36.9	34.2
Phenylalanine	22.4	28.9
Cystine and cysteine	Not determined	Not determined

					5					10
Be	(*)	Leu	Thr	Ile	Glx^+	Ser	(*)	Pro	Phe	Asx^+
Na 1–3	Lys	Leu	Thr	Ile	Glu	Ser	Thr	Pro	Phe	Asn
					15					20
Be	Val	Ala	Glx^+	Gly	(*)	Glx^+	Val	Leu	Leu	
Na 1–3	Val	Ala	Glu	Gly	Lys	Glu	Val	Leu	Leu	Leu
					24					
Na 1–3	Val	His	Asn	Leu						

FIG. 2. Amino-terminal sequences of CEA preparations Be and Na 1–3. Residues were identified by gas chromatography and confirmed by amino acid analysis. (*), amino acid could not be defined, $^+$, identified only after hydrolysis, therefore dicarboxylic acids cannot be distinguished from amide forms.

lished. The sequence of Na 1–3 was determined for 24 positions (Fig. 2) and was shown to be identical to that of Be. In addition, however, it was possible to establish that the residues at positions 1 and 15 were lysine, that at position 7 was threonine, and to establish unequivocally the residues at positions 5, 10, 13, and 16.

Interpretation of the sequence analysis depends upon the yield of amino acid in one of the first few positions. The first amino acid, lysine, is not recovered quantitatively and therefore yields were calculated for leucine at position 2. Assuming that CEA has a molecular weight of 200,000, a protein content of 50%, and consists of a single polypeptide chain, the yield for preparation Be was 35% of theoretical, and that for Na 1–3 was 70% of theoretical. Yield differences may reflect inherent differences in the two CEA preparations, or may be due to the additional purification of Na 1–3. Whatever the explanation, this range of yields is consistent with that obtained in our sequencer with other homogeneous glycoproteins (e.g., immunoglobulin heavy chains). The yield data indicate that the polypeptide chain being sequenced accounts for all or almost all of the protein with unblocked amino-terminal amino acids in these preparations. The single sequence obtained demonstrates that the unblocked polypeptide chains are homogeneous for the first 24 residues. Additional studies will be required to determine whether these preparations contain significant amounts of protein with blocked amino-terminal residues.

The finding of identical sequences in the CEA preparations derived from two different tumors suggests that the polypeptide chain being sequenced will be a constant constituent of material with CEA activity. It has not, however, been proven that the proteins being sequenced are indeed associated with CEA activity. It is possible, for example, that these proteins are of the same size and have the same electrophoretic characteristics as the material carrying CEA activity, but that the activity is associated with a different molecular species. Additional studies are presently in progress in an attempt to establish firm identity between this protein and the molecules carrying CEA antigenic activity.

We are, at the present time, isolating quantities of CEA glycoprotein to enable determination of the complete amino acid sequence of the relevant protein when the various problems described above are solved. This should be of great interest, even if the carbohydrate moiety contains the antigenic activity, since the structural information will permit comparisons to be made with other glycoproteins extracted from the membranes of malignant, fetal, and normal adult cells.

REFERENCES

1. Gold, P. 1971. Antigenic reversion in human cancer. *Annu. Rev. Med.* **22:**85.
2. Gold, P., and S. O. Freedman. 1965. Specific carcinoembryonic antigens of the human digestive system. *J. Exp. Med.* **122:**467.
3. Coligan, J. E., J. T. Lautenschleger, M. L. Egan, and C. W. Todd. 1972. Isolation and characterization of carcinoembryonic antigen. *Immunochemistry.* **9:**377.
4. Martin, F., and M. S. Martin. 1970. Distribution of the carcinoembryonic antigen in the human digestive system. *Int. J. Cancer.* **6:**352.
5. Krupey, J., P. Gold, and S. O. Freedman. 1968. Physicochemical studies of the carcinoembryonic antigens of the human digestive system. *J. Exp. Med.* **128:** 387.
6. Egan, M. L., J. T. Lautenschleger, J. E. Coligan, and C. W. Todd. 1972. Radioimmune assay of carcinoembryonic antigen. *Immunochemistry.* **9:**289.
7. Davis, B. J. 1964. Disc electrophoresis. II. Method and application to human serum proteins. *Ann. N. Y. Acad. Sci.* **121:**404.
8. Weber, K., J. R. Pringle, and M. Osborn. Measurement of molecular weights by electrophoresis on SDS-acrylamide gels. *In* Methods in Enzymology. S. P. Colowick and N. O. Kaplan, editors. Academic Press, Inc., New York. 2nd edition. In press. **11.**
9. Zacharius, R. M., T. E. Zell, J. H. Morrison, and J. J. Woodlock. 1969. Glycoprotein staining following electrophoresis on acrylamide gels. *Anal. Biochem.* **30:**148.
10. Pisano, J., and T. Bronzert. 1969. Analysis of amino acid phenylthiohydantoins by gas chromatography. *J. Biol. Chem.* **244:**5597.
11. Kaplan, A. P., L. E. Hood, W. D. Terry, and H. Metzger. 1971. Amino-terminal sequences of human immunoglobulin heavy chains. *Immunochemistry.* **8:**801.